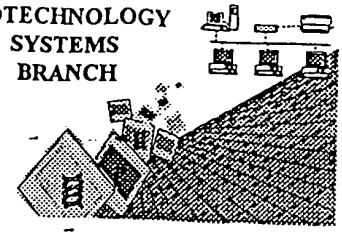
ep

BIOTECHNOLOGY
SYSTEMS
BRANCH



RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/643,752
Source: OIPE
Date Processed by STIC: 8-25-03

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

1) **INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,**
2) **TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY**

FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 4.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:
http://www.uspto.gov/web/offices/pac/checker

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.
Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.
Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<http://www.uspto.gov/ebc/efs/downloads/documents.htm>, **EFS Submission User Manual** - ePAVE)
2. **U.S. Postal Service:** Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. **Hand Carry directly to:**
   U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7th Floor, Examiner Name, Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202
   Or
   U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. **Federal Express, United Parcel Service, or other delivery service to:** U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 04/24/2003

Raw Sequence Listing Error Summary

ERROR DETECTED | **SUGGESTED CORRECTION** | **SERIAL NUMBER:** 10/643,752

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1	Wrapped Nucleics Wrapped Aminos	The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
2	Invalid Line Length	The rules require that a line not exceed 72 characters in length. This includes white spaces.
3	Misaligned Amino Numbering	The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
4	Non-ASCII	The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
5	Variable Length	Sequence(s)____contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
6	PatentIn 2.0 "bug"	A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
7	Skipped Sequences (OLD RULES)	Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence: (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) ~~(i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)~~ (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) This sequence is intentionally skipped Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
8	Skipped Sequences (NEW RULES)	Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence. <210> sequence id number <400> sequence id number 000
9	Use of n's or Xaa's (NEW RULES)	Use of n's and/or Xaa's have been detected in the Sequence Listing. Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present. In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
10	Invalid <213> Response	Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
11 X	Use of <220>	Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section. (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
12	PatentIn 2.0 "bug"	Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
13	Misuse of n	n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.

AMC/MH – Biotechnology Systems Branch – 08/21/2001

OIPE

RAW SEQUENCE LISTING DATE: 08/25/2003
PATENT APPLICATION: **US/10/643,752** TIME: 14:10:45

Input Set : **A:\LS5-001.ST25.txt**
Output Set: **N:\CRF4\08252003\J643752.raw**

```
  3 <110> APPLICANT: Liu, David R.
  4       Gartner , Zev J.
  5       Doyon, Jeffrey B.
  6       Calderone , Christopher T.
  7       Kanan, Matthew W.
  8       Li, Xiaoyu
  9       Snyder, Thomas M.
 10       Rosenbaum, Daniel M.
 12 <120> TITLE OF INVENTION: Evolving New Molecular Function
 14 <130> FILE REFERENCE: LS5-001
C--> 16 <140> CURRENT APPLICATION NUMBER: US/10/643,752
C--> 16 <141> CURRENT FILING DATE: 2003-08-19
 16 <150> PRIOR APPLICATION NUMBER: US 60/404,395
 17 <151> PRIOR FILING DATE: 2002-08-19
 19 <150> PRIOR APPLICATION NUMBER: US 60/419,667
 20 <151> PRIOR FILING DATE: 2002-10-18
 22 <150> PRIOR APPLICATION NUMBER: US 60/432,812
 23 <151> PRIOR FILING DATE: 2002-12-11
 25 <150> PRIOR APPLICATION NUMBER: US 60/444,770
 26 <151> PRIOR FILING DATE: 2003-02-04
 28 <150> PRIOR APPLICATION NUMBER: US 60/457,789
 29 <151> PRIOR FILING DATE: 2003-03-26
 31 <150> PRIOR APPLICATION NUMBER: US 60/469,866
 32 <151> PRIOR FILING DATE: 2003-05-12
 34 <150> PRIOR APPLICATION NUMBER: US 60/479,494
 35 <151> PRIOR FILING DATE: 2003-06-18
 37 <160> NUMBER OF SEQ ID NOS: 125
 39 <170> SOFTWARE: PatentIn version 3.1
 41 <210> SEQ ID NO: 1
 42 <211> LENGTH: 64
 43 <212> TYPE: DNA
 44 <213> ORGANISM: Artificial Sequence
 46 <220> FEATURE:
 47 <223> OTHER INFORMATION: Template Encoding Parent Molecule 1
 49 <400> SEQUENCE: 1
 50 cgagcagcac cagcgcactc cgcctggatc cgccccgggt gcacgcgact cctacgggct     60
 52 ccaa                                                                64
 55 <210> SEQ ID NO: 2
 56 <211> LENGTH: 64
 57 <212> TYPE: DNA
 58 <213> ORGANISM: Artificial Sequence
 60 <220> FEATURE:
 61 <223> OTHER INFORMATION: Template Encoding Parent Molecule 2
```

Does Not Comply
Corrected Diskette Needed

P. 2-6

RAW SEQUENCE LISTING DATE: 08/25/2003
PATENT APPLICATION: **US/10/643,752** TIME: 14:10:45

Input Set : **A:\LS5-001.ST25.txt**
Output Set: **N:\CRF4\08252003\J643752.raw**

63 <400> SEQUENCE: 2
64 cgagcagcac cagcgagtcc cgcctgggga tgccccgggt gggcgcgact ccaacgggct 60
66 ccaa 64
69 <210> SEQ ID NO: 3
70 <211> LENGTH: 64
71 <212> TYPE: DNA
72 <213> ORGANISM: Artificial Sequence
74 <220> FEATURE:
75 <223> OTHER INFORMATION: Recombined Daughter Template
77 <400> SEQUENCE: 3
78 cgagcagcac cagcgcactc cgcctgggga tgccccgggt gggcgcgact cctacgggct 60
80 ccaa 64
83 <210> SEQ ID NO: 4
84 <211> LENGTH: 64
85 <212> TYPE: DNA
86 <213> ORGANISM: Artificial Sequence
88 <220> FEATURE:
89 <223> OTHER INFORMATION: Recombined Daughter Template
91 <400> SEQUENCE: 4
92 cgagcagcac cagcgagtcc cgcctggatc cgccccgggt gcacgcgact ccaacgggct 60
94 ccaa 64
97 <210> SEQ ID NO: 5
98 <211> LENGTH: 10
99 <212> TYPE: DNA
100 <213> ORGANISM: Artificial Sequence
102 <220> FEATURE:
103 <223> OTHER INFORMATION: Reagent
105 <400> SEQUENCE: 5
106 aattcgtacc 10
109 <210> SEQ ID NO: 6
110 <211> LENGTH: 11
111 <212> TYPE: DNA
112 <213> ORGANISM: Artificial Sequence
114 <220> FEATURE:
115 <223> OTHER INFORMATION: Template E
117 <400> SEQUENCE: 6
118 tggtacgaat t 11
121 <210> SEQ ID NO: 7
122 <211> LENGTH: 31
123 <212> TYPE: DNA
124 <213> ORGANISM: Artificial Sequence
126 <220> FEATURE:
127 <223> OTHER INFORMATION: Template H
129 <400> SEQUENCE: 7
130 tcgcgagcgt acgctcgcga tggtacgaat t 31
133 <210> SEQ ID NO: 8
134 <211> LENGTH: 20
135 <212> TYPE: DNA
136 <213> ORGANISM: Artificial Sequence

Insufficient response
Please give source of
genetic material.
(see item 11 on error
summary report.

RAW SEQUENCE LISTING DATE: 08/25/2003
PATENT APPLICATION: **US/10/643,752** TIME: 14:10:45

Input Set : **A:\LS5-001.ST25.txt**
Output Set: **N:\CRF4\08252003\J643752.raw**

```
138 <220> FEATURE:
139 <223> OTHER INFORMATION: Template
141 <400> SEQUENCE: 8
142 tggtacgaat tcgactcggg                                              20
145 <210> SEQ ID NO: 9
146 <211> LENGTH: 10
147 <212> TYPE: DNA
148 <213> ORGANISM: Artificial Sequence
150 <220> FEATURE:
151 <223> OTHER INFORMATION: Reagent
153 <400> SEQUENCE: 9
154 cccgagtcga                                                         10
157 <210> SEQ ID NO: 10
158 <211> LENGTH: 50
159 <212> TYPE: DNA
160 <213> ORGANISM: Artificial Sequence
162 <220> FEATURE:
163 <223> OTHER INFORMATION: Template
165 <400> SEQUENCE: 10
166 tggtgcggag ccgccgtgac gggtgatacc acctccgagc cgaggagccg              50
169 <210> SEQ ID NO: 11
170 <211> LENGTH: 50
171 <212> TYPE: DNA
172 <213> ORGANISM: Artificial Sequence
174 <220> FEATURE:
175 <223> OTHER INFORMATION: Template
177 <220> FEATURE:
178 <221> NAME/KEY: misc_feature
179 <222> LOCATION: (17)..(17)
180 <223> OTHER INFORMATION: N is A, C, T or G
183 <220> FEATURE:
184 <221> NAME/KEY: misc_feature
185 <222> LOCATION: (19)..(19)
186 <223> OTHER INFORMATION: N is A, C, T or G
189 <220> FEATURE:
190 <221> NAME/KEY: misc_feature
191 <222> LOCATION: (21)..(21)
192 <223> OTHER INFORMATION: N is A, C, T or G
195 <220> FEATURE:
196 <221> NAME/KEY: misc_feature
197 <222> LOCATION: (23)..(24)
198 <223> OTHER INFORMATION: N is A, C, T or G
201 <400> SEQUENCE: 11
W--> 202 tggtgcggag ccgccgncna ncnngatacc acctccgagc cgaggagccg          50
205 <210> SEQ ID NO: 12
206 <211> LENGTH: 10
207 <212> TYPE: DNA
208 <213> ORGANISM: Artificial Sequence
210 <220> FEATURE:
```

RAW SEQUENCE LISTING DATE: 08/25/2003
PATENT APPLICATION: **US/10/643,752** TIME: 14:10:45

Input Set : **A:\LS5-001.ST25.txt**
Output Set: **N:\CRF4\08252003\J643752.raw**

211 <223> OTHER INFORMATION: Reagent
213 <400> SEQUENCE: 12
214 cacccgtcac 10
217 <210> SEQ ID NO: 13
218 <211> LENGTH: 10
219 <212> TYPE: DNA
220 <213> ORGANISM: Artificial Sequence
222 <220> FEATURE:
223 <223> OTHER INFORMATION: Reagent
225 <220> FEATURE:
226 <221> NAME/KEY: misc_feature
227 <222> LOCATION: (2)..(3)
228 <223> OTHER INFORMATION: N is A, T, C or G
231 <220> FEATURE:
232 <221> NAME/KEY: misc_feature
233 <222> LOCATION: (5)..(5)
234 <223> OTHER INFORMATION: N is A, T, C or G
237 <220> FEATURE:
238 <221> NAME/KEY: misc_feature
239 <222> LOCATION: (7)..(7)
240 <223> OTHER INFORMATION: N is A, T, C or G
243 <220> FEATURE:
244 <221> NAME/KEY: misc_feature
245 <222> LOCATION: (9)..(9)
246 <223> OTHER INFORMATION: N is A, T, C or G
249 <400> SEQUENCE: 13
W--> 250 **cnngntngnc** **10**
253 <210> SEQ ID NO: 14
254 <211> LENGTH: 11
255 <212> TYPE: DNA
256 <213> ORGANISM: Artificial Sequence
258 <220> FEATURE:
259 <223> OTHER INFORMATION: Template 1a-1c
261 <400> SEQUENCE: 14
262 tggtacgaat t 11
265 <210> SEQ ID NO: 15
266 <211> LENGTH: 17
267 <212> TYPE: DNA
268 <213> ORGANISM: Artificial Sequence
270 <220> FEATURE:
271 <223> OTHER INFORMATION: Template 2a-2c
273 <400> SEQUENCE: 15
274 ttaacgagag atagtct 17
277 <210> SEQ ID NO: 16
278 <211> LENGTH: 23
279 <212> TYPE: DNA
280 <213> ORGANISM: Artificial Sequence
282 <220> FEATURE:
283 <223> OTHER INFORMATION: Template 3a-3c

RAW SEQUENCE LISTING DATE: 08/25/2003
PATENT APPLICATION: **US/10/643,752** TIME: 14:10:45

Input Set : **A:\LS5-001.ST25.txt**
Output Set: **N:\CRF4\08252003\J643752.raw**

285 <400> SEQUENCE: 16
286 tatctacaga gtagtctaat gac 23
289 <210> SEQ ID NO: 17
290 <211> LENGTH: 14
291 <212> TYPE: DNA
292 <213> ORGANISM: Artificial Sequence
294 <220> FEATURE:
295 <223> OTHER INFORMATION: Reagent 4a-4c
297 <400> SEQUENCE: 17
298 cagcaattcg tacc 14
301 <210> SEQ ID NO: 18
302 <211> LENGTH: 16
303 <212> TYPE: DNA
304 <213> ORGANISM: Artificial Sequence
306 <220> FEATURE:
307 <223> OTHER INFORMATION: Reagent 5a-5c
309 <400> SEQUENCE: 18
310 ctcagctctc tcgtta 16
313 <210> SEQ ID NO: 19
314 <211> LENGTH: 18
315 <212> TYPE: DNA
316 <213> ORGANISM: Artificial Sequence
318 <220> FEATURE:
319 <223> OTHER INFORMATION: Reagent 6a-6c
321 <400> SEQUENCE: 19
322 ggctcagcct ctgtagat 18
325 <210> SEQ ID NO: 20
326 <211> LENGTH: 11
327 <212> TYPE: DNA
328 <213> ORGANISM: Artificial Sequence
330 <220> FEATURE:
331 <223> OTHER INFORMATION: Template 15
333 <400> SEQUENCE: 20
334 tatagatcag c 11
337 <210> SEQ ID NO: 21
338 <211> LENGTH: 11
339 <212> TYPE: DNA
340 <213> ORGANISM: Artificial Sequence
342 <220> FEATURE:
343 <223> OTHER INFORMATION: Template 17
345 <400> SEQUENCE: 21
346 ttaacgagag a 11
349 <210> SEQ ID NO: 22
350 <211> LENGTH: 11
351 <212> TYPE: DNA
352 <213> ORGANISM: Artificial Sequence
354 <220> FEATURE:
355 <223> OTHER INFORMATION: Template 18
357 <400> SEQUENCE: 22

The type of errors shown exist throughout the Sequence Listing. Please check subsequent sequences for similar errors.

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 08/25/2003
PATENT APPLICATION: **US/10/643,752** TIME: 14:10:46

Input Set : **A:\LS5-001.ST25.txt**
Output Set: **N:\CRF4\08252003\J643752.raw**

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:11; N Pos. 17,19,21,23,24
Seq#:13; N Pos. 2,3,5,7,9
Seq#:31; N Pos. 11,12,13,14,15,16,22,23,24,25,26,27,33,34,35,36,37,38
Seq#:34; N Pos. 17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36
Seq#:60; N Pos. 14

VERIFICATION SUMMARY DATE: 08/25/2003
PATENT APPLICATION: **US/10/643,752** TIME: 14:10:46

Input Set : **A:\LS5-001.ST25.txt**
Output Set: **N:\CRF4\08252003\J643752.raw**

L:16 M:270 C: Current Application Number differs, Replaced Current Application No
L:16 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:202 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11 after pos.:0
L:250 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13 after pos.:0
L:484 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:31 after pos.:0
L:526 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:34 after pos.:0
L:844 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:60 after pos.:0